Forschungsverbund
Ost- und Südosteuropa

Transition Economies

Are Transition Countries Overbanked? The Effect of Institutions on Bank Market Entry

Christa Hainz

General Equilibrium Model of an Economy with a Futures Market

Roman Cech

forost Arbeitspapier Nr. 15
September 2003

Vorwort

Mit dem hier vorgelegten Band der Veröffentlichungen von *forost* legt die Forschungsgruppe I zwei Arbeiten vor, die sich mit speziellen Problemen der Transformation unter Anwendung modelltheoretischer Ansätze auseinander setzen.

Roman Cech befasst sich mit der Frage, wie sich in einer Wirtschaft, die über Zukunftsmärkte verfügt und deren Akteure unterschiedliche Risikoaversionen besitzen, ein Gleichgewicht herausbildet. Dabei sind seine *Ergebnisse*, dass die Zukunftsmärkte die Entwicklung der realen Wirtschaft beeinflussen und auf die Einkommensverteilung in einer Volkswirtschaft Einfluss nehmen für die Entwicklung der Transformationsgesellschaften und ihre Annäherung an die EU von besonderer Bedeutung.

Christa Hainz überprüft die These, dass die Transformationsstaaten „overbanked" seien, und sucht nach dem Grund für dieses Phänomen. Dabei zeigt sie, dass ein entscheidender Grund in den noch insgesamt unzureichend funktionierenden Institutionen während der Anfangphase der Transformation zu suchen ist. Daher ist zu erwarten, dass mit der Festigung der Institutionen auch eine Konsolidierung des Bankensystems einhergehen wird. Daher muss in allen Transformationsstaaten auch unter dem Gesichtspunkt der Konsolidierung des Bankensystems – was eine wichtige Voraussetzung für die weitere Entwicklung der Transformationsstaaten ist – der Festigung der Institutionen besondere Aufmerksamkeit gewidmet werden.

Beide Arbeiten sind von grundsätzlicher Bedeutung für der Annäherung der Transformationsstaaten an die EU und den in diesem Prozess zu erwartenden Entwicklungen, und sie erlauben Rückschlüsse auf die notwendigen wirtschaftspolitischen Maßnahmen.

München, im August 2003 — Hermann Clement forost

INHALT

forost

Are Transition Countries Overbanked? The Effect of Institutions on Bank Market Entry

Christa Hainz*

April, 2003

Abstract

The popular notion that transition countries are overbanked is challenged here. We study the decision for market entry and the optimal number of banks in a Salop-model. We show that the amount of collateral depends on the distance between bank and firm as well as the quality of the institutional environment. We analyze how the number of banks decreases as the institutional environment improves. Moreover, we find that market entry is insufficient because new entrants do not consider the positive effects of their entry decision on social welfare sufficiently, i.e. the reduction of collateral requirement and the increase in the average liquidation value.

JEL-Classification: D43, G21, G34, L13, P31, P34

Keywords: Transition economies, bank competition, market entry, corporate finance, corporate governance

---

*Department of Economics, University of Munich, Akademiestr. 1/III, 80799 Munich, Tel.: (49 89) 2180 3232, Fax.: (49 89) 2180 2767, e-mail: christa.hainz@lrz.uni-muenchen.de. The author would like to thank Dilip Mookherjee, Ernst Maug, Kira Börner, Roman Cech, Andrea Gyulai-Schmidt, Isabelle Kronawitter, Richard Schmidtke and Monika Schnitzer, as well as seminar participants at the Austrian National Bank, the ESEM 2002, the EARIE Meeting 2002, the German Economic Association Meeting 2002 and the CESifo Conference on The Economics of Organisation and Corporate Governance Structures for helpful comments and suggestions. The usual disclaimer applies. Financial support by FOROST trough project "On the role of banks for corporate finance and restructuring in transition economies" is gratefully acknowledged.

1. Introduction

The number of banks in transition countries soared in the beginning of the transition process. Russia is the most popular example: The number of banks peaked in 1996 at about 2,000, but it sunk significantly after the crisis in 1998 to about 1,300 in 2000 (EBRD, 2000; OECD, 2000). The strong decrease in the number of banks suggests that Russia was overbanked during the 1990s. Russia can be used to illustrate the situation in Eastern Europe where the banking sector is often described as overbanked but underserviced (Bonin et al., 1998). Which are the factors that render market entry so attractive?

To answer the questions on market entry we set up a theoretical model of spatial competition where banks have to bear the transportation costs. In transition countries banks collateralize nearly all credit contracts (Fan et al., 1996; Bratkowski et al., 2000). However, banks differ in their liquidation values of collateralized assets. Thus, the model provides new insights into the behavior of banks in transition countries. It also contributes to the theoretical literature on spatial competition. A traditional Salop-model with mill pricing shows that too many firms (or banks) enter.[1] In contrast, we use a more realistic model with delivered pricing since banks face a loss if collateralized assets are liquidated. The surprising result is that the economy is underbanked.

Due to the deficient legal and institutional environment in transition countries, the model has to portray the problems associated with collateralization. This field of problems is taken into account as follows: Banks have higher liquidation values of collateral if they are located closer to a firm. It is shown that collateralization is used to solve the moral hazard problem of finance. A bank makes a positive profit if it has a comparative advantage in liquidating collateralized assets of any particular firm. As banks cannot price-discriminate, they offer the same contract to all borrowers.

However, this analysis shows that market entry has interesting implications for the industry as a whole. As we would expect, market entry reduces the market share of each individual bank. At the same time, the amount of collateral and the repayment decrease as the distance between the firm and the bank is reduced. This is the negative externality which is typically found in a Salop-model. Interestingly, market entry increases the liquidation payoff, which is obtained from the assets of the marginal borrower. Therefore, it achieves a higher average return if assets have to be liquidated. However, the negative effects of market entry on bank profit dominate. And therefore the profit of each bank decreases as a new bank enters. From a social welfare perspective we find that the number of banks in equilibrium is too low. The banks do not internalize the positive effects of entry on social welfare. Market entry increases social welfare by reducing the costs of collateralization in two ways: first, by lowering the amount of collateral and second, by increasing the average liquidation value. Due to these positive externalities, the equilibrium number of banks is lower than the socially optimal number.

This paper is related to two areas in the literature: banking in transition countries and spatial competition. The empirical studies propose that transition countries are overbanked but underserviced (Bonin and Wachtel, 1999). Jaffee and Levonian (2001)

[1] The distinction between delivered and mill pricing is made by Gabszewicz and Thisse (1986).

carry out an empirical analysis of the number of banks in transition economies. They calculate a benchmark value for the number of banks in transition countries in 1995. To calculate the benchmark they use data on the number of banks in OECD countries in order to find out the main determinants of the structural characteristics of the banking system. In their calculation of the number of banks, GDP is the only significant variable.[2] However, the explanatory power is limited as the $R^2 = 0.617$ suggests. The authors conclude that other factors, such as regulation, have a strong influence on the number of banks. With regards to transition economies, they show that the number of banks is lower than the benchmark value in the Czech Republic, Hungary and Poland, whereas in all other countries it is higher.

The empirical literature on banking in transition economies also explores the effects of a deficient legal and institutional environment. McNulty and Harper (2001) find in their regression that these deficiencies are responsible for the low degree of financial intermediation. But how does the poor legal and institutional environment affect the banking sector? This relationship becomes most obvious under extremely widespread collateralization (Fan et al., 1996; Bratkowski et al., 2000). First, if banks rely on public contract enforcement they encounter several difficulties. In Russia, for instance, the bailiffs face an overwhelming caseload. Therefore, they carefully select the cases that they solve. As they usually have to travel to the location of a judgment creditor by public transportation, the incentive to solve cases in distant locations is low (Kahn, 2002). Second, as public contract enforcement often works insufficiently, some banks may prefer private contract enforcement (McMillan and Woodruff, 2000). This means that they have to bear significant costs for contract enforcement; moreover, they might have different costs. Finally, secondary markets do not function perfectly. Due to generally bad economic conditions, the demand for the assets is low and so are the returns for the bank. Since the institutional environment is poor, agents which have superior information or are integrated into networks can achieve a higher return.[3] The popularity of collateralization is surprising as the liquidation payoff for the bank is low. Nevertheless, collateralization helps solve the incentive problems associated with debt financing: adverse selection, moral hazard and state verification (Bester, 1985; Holmström, 1996; Bester, 1994).

Thus far, the theoretical literature on banking in transition countries has not studied the optimal number of banks intensively; it has largely been focused on non-performing loans and competition. However, the intensity of bank competition and the number of banks are related to each other. Schnitzer (1999) studies the bank's incentive to invest in perfect screening in a Salop-model. The analysis reveals that the incentive to invest in screening depends on the number of uninformed competitors. The screening costs determine whether all banks screen or no bank screens, and they also determine the extent of market entry. The first best number of banks enter only if screening costs are high. There are too many banks present in the case of low and intermediate screening costs. Generally, banks either screen too much from a social welfare point of view (if

[2] All other variables, i.e. population, size of the country, gross saving ratio, and ratio of non-resident claims to total claims on banks, are insignificant.

[3] Koford and Tschoegl (1999) provide evidence from Bulgaria.

screening costs are low) or they do not screen at all (if screening costs are intermediate).[4] Solely in the latter case is the banking sector overbanked but underserviced - as described in the transition literature.

The second area in the literature which this paper is complementary to is spatial competition. In his seminal paper, Salop (1979) develops a spatial model of market entry with mill pricing. From a normative point of view there are too many firms in the market. The case of a uniform delivered price is studied by Gromberg and Meyer (1981). They distinguish between extreme price competition and collusion among firms. However, they do not consider social welfare. To our knowledge, only the paper by Matsumura (2000) shows that too few firms enter. He considers the integer problem of the number of firms and shows that excess entry theorem occurs if the marginal production cost is constant. Whereas if the marginal production cost is increasing, the excess-entry theorem no longer holds true.

The paper is organized as follows: In Section 2, the Salop model and the optimal credit contracts are studied. In Section 3, we first analyze market entry into the banking sector and then compare the number of banks in equilibrium with the socially optimal number. Finally, in Section 4, the results are discussed by relating them to the empirical evidence.

2. Bank Competition and Asset Specificity

2.1. Model

In this model firms which want to undertake an investment project with costs of I are studied. Each firm has an asset endowment of A. Firms can only realize the project if they receive a credit because they do not have sufficient liquid means. The expected return of the project depends on the effort of the firm's manager. If effort E is exerted, the probability for the successful outcome with a high return of X increases from p_L to p_H. In the case of failure, the return is 0. Firms are uniformly distributed along a circular road of length 1 and their total mass is normalized to 1.

In a first best environment, the bank observes the effort level of the manager and, as we assume that $p_H X - E > I > p_L X$, the contract determines that the manager exerts effort E. However, as the effort is not observable in reality, the credit contract has to be designed in a manner that gives the firm's manager an incentive to exert effort. The manager's incentive depends on the state of the world where the return is determined by the gross return and the repayment to the bank. The firm has to repay R in the case of success. If the project fails, the amount of collateral L is liquidated by the bank. It is assumed that the asset endowment is high enough to avoid problems of insufficient collateralizable wealth, i.e. $A \geq \frac{E}{p_H - p_L} p_L$.[5]

[4] A different set up is chosen by Broecker (1990): He studies a model with imperfect but costless screening. It is shown that the higher the number of banks, the higher is the repayment as the average quality of the applying firm decreases.

[5] A monopolistic bank demands collateral of $\frac{E}{p_H - p_L} p_L$ (see Hainz, 2003). Holmström (1996) studies

The banking sector consists of N banks. Each bank is allowed to locate in only one location. The banks do not choose their location, they are automatically located equidistantly from one another. In this way maximal differentiation between banks is exogenously imposed. The banks compete in Bertrand fashion for firms which are asking for credit.[6] Banks as well as firms are risk-neutral. Due to the substantial costs in the case of liquidation, the return to the bank, denoted by α, is lower than the continuation value within the firm, i.e. $0 \leq \alpha \leq 1$. In transition economies, liquidation leads to a considerable discount because public as well as private contract enforcement is expensive and secondary markets function only very imperfectly.

The time structure of the game is summarized in the following figure:

Figure 1: Time Structure

In the following section, we study the design of a credit contract for a given number of competitors in the banking sector. In the next step, the number of banks will be determined endogenously.

2.2. Bank Competition and Collateralized Credit Contracts

The bank's liquidation payoffs of the collateralized assets depend on the distance between the location of the bank and the firm. Thus, α is higher for firms in close proximity to a bank. To keep the analysis as simple as possible, we do not consider transaction costs and set $\alpha = 1$ for firms located at the same place as a bank. The further the distance which a firm has to travel to a bank, the lower is α. It is assumed that α decreases proportionally with the distance travelled. The liquidation value is lowest for a firm that is located directly opposite the bank on the circle, and is denoted by $\underline{\alpha}$. The liquidation value of the marginal firm is denoted by α^K.

There are two possible interpretations of what termed "distance" in our model. First, if there is spatial competition, it is the physical distance that causes costs.[7] For the bank closest to a firm the costs of enforcing the contract are lower than that of any other competitor. Second, banks specialize in financing firms from certain sectors. In that case,

the problems which are caused if firms do not possess sufficient collateral.

[6] The costs of funds are normalized to zero. In this analysis we focus on bank competition on credit markets. For transition countries Dittus and Prowse (1996) show that only a few banks take deposits, which they transfer to the credit-granting banks through the money market.

[7] There is evidence from Belgium for spatial competition in the credit market (Degryse and Ongena, 2002).

the bank that is perfectly specialized in the business of the firm has the best expertise in liquidating the assets because it knows the market and its players best. If they finance firms that are not belonging to their core group of customers, banks have higher costs translating into a lower liquidation value. This is particularly important in transition economies since the secondary markets work poorly and therefore the expertise of each individual bank plays a crucial role.

Furthermore, we assume that the bank cannot observe the location of each individual firm that is asking for credit and that firms ask for credit at the closest bank if the banks offer identical credit contracts. This assumption rules out price discrimination. Although the relative size of the different effects changes, price discrimination would not change the insights gained from this analysis. The assumptions can be justified as follows: In the case of spatial competition, public contract enforcement renders predicting the liquidation value of a particular asset difficult, as the example Russia shows. The incentives of the bailiffs to solve a case depend on the various features of a case, one of which is distance. Their incentives to solve a particular case should decrease the further away the judgment creditor is located. However, there might be economics of scale if various cases have to be solved at a particular location. The banks knowledge is limited to information pertaining to their own customers. Hence, firms and bailiffs should be better informed than banks about disputes of other firms in their surrounding area and therefore a more precise assessment of the liquidation value. Moreover, banks that enter the market do not know the markets for collateralized assets very well. Firms are better informed than banks with respect to the value of the collateralized asset in the case of failure.

The objective of the representative bank i is to maximize profit Π_i^B given by:

$$\max_{R,L} \ \Pi_i^B = \frac{1}{N}\left(\int_{\alpha^K}^{1} (p_H R + (1-p_H)\,\alpha L - I)\, d\alpha\right) \tag{2.1}$$

The bank's profit depends on its market share, given by $\frac{1}{N}$, and its profit from the firms financed, given by $\int_{\alpha^K}^{1} (p_H R + (1-p_H)\,\alpha L - I)\, d\alpha$. Customers are located on the right and the left hand side of the bank. The bank serves all firms located on the circle between the marginal firm, having a liquidation value of α^K, and the firm at its own location, having a liquidation value of 1. The liquidation value for the marginal firm, $\alpha^K = \left(1 - \frac{(1-\underline{\alpha})}{N}\right)$, depends on the lowest liquidation value and the number of banks. The latter determines the distance between the marginal firm and the bank. The liquidation value is lowest, i.e. $\alpha = \underline{\alpha}$, if a firm is financed by the bank that is located opposite on the circle.

When determining the terms of the contract $\{R, L\}$ the bank has to take into account the following constraints: First, the credit contract has to provide sufficient incentives to the firm's manager to exert effort; failing which, the bank's expected profit is negative (**?). The firm's incentive compatibility constraint is defined by:

$$\begin{aligned} & p_H (A + X - R) + (1-p_H)(A - L) - E \\ \geq \ & p_L (A + X - R) + (1-p_L)(A - L) \end{aligned} \tag{IC-F}$$

$$R \leq X + L - \frac{E}{\Delta p}$$

with $\Delta p = (p_H - p_L)$. We assume that $0 > p_H X - I - \frac{E p_H}{\Delta p} \geq -\frac{E p_L}{\Delta p}\left(p_H + (1 - p_H)\,\alpha^K\right)$, as (a) it is not possible to solve the problem without collateralization and (b) the payoff of investment is high enough to cover the costs associated with collateralization.[8]

Second, the firm has to be at least as well off with the credit financed investment as without it. The firm's participation constraint is given by:

$$p_H\,(A + X - R) + (1 - p_H)\,(A - L) - E \geq A \tag{PC-F}$$

And finally, the bank has to consider the offers of the competing banks. The credit contract, which results from the maximizing the bank's profit function, subject to the above-mentioned constraints, is described in Proposition 1:

Proposition 1. *All banks offer the same credit contract which determines the repayment and the amount of collateral, i.e.* $R = \frac{\alpha^K (1-p_H)\left(X - \frac{E}{\Delta p}\right) + I}{p_H + (1-p_H)\alpha^K}$ *and* $L = \frac{-p_H X + \frac{E}{\Delta p} p_H + I}{p_H + (1-p_H)\alpha^K}$.

Proof: See Appendix.

By demanding a collateral the firm's payoff in the case of failure is reduced, the difference between state-contingent payoffs increases and thus the incentive of the firm's manager to exert effort increases. As (IC-F) shows, the collateral requirement L also influences the repayment R, which can be demanded without destroying the manager's incentive. The amount of collateral is determined by the liquidation value of the marginal firm's assets, α^K. The marginal firm is located half-way between two banks. For this firm, the banks engage in perfect competition as their liquidation value assessment for this firm is the same. The marginal firm has to generate an expected profit of zero for each bank, i.e. $p_H\left(X + L - \frac{E}{\Delta p}\right) + (1 - p_H)\,\alpha^K L - I = 0$. As it is not possible to discriminate between firms from different locations, the bank demands the same collateral from all firms. The assets of the other firms, which are located closer to the bank, have a higher liquidation value, and therefore these firms yield a positive profit for the bank.[9]

Offering the credit contract described in Proposition 1, the bank's profit is:

$$\Pi_i^B = \frac{1}{N}\left(p_H R + (1 - p_H)\left(\frac{1 + \alpha^K}{2}\right) L - I\right) \tag{2.2}$$

inserting $R = X + L - \frac{E}{\Delta p}$ and $L = \frac{-p_H X + \frac{E}{\Delta p} p_H + I}{p_H + (1-p_H)\alpha^K}$ yields

$$\Pi_i^B = \frac{1}{N}\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) \frac{(1 - p_H)\left(1 - \alpha^K\right)}{2\left(p_H + (1 - p_H)\,\alpha^K\right)}$$

[8] Formally, the first assumption implies that the bank's zero profit constraint after inserting (PC-F) would not hold if $L = 0$. The second assumption affirms that the firm's participation constraint holds for the optimal L determined in Proposition 1.

[9] If the banks knew where a firm is located, the collateral would be higher for all firms except the marginal firm. The explanation is as follows: Bank B which is further away from the firm is less specialized than bank A which is closest. Thus, bank B needs a higher repayment to make zero profit. Accordingly, bank B can demand a higher repayment than in the case without price discrimination, and it can extract more rent.

3. Market entry into the banking sector

3.1. Individual Entry Decision

So far the number of banks has been taken as given. What does a change in the number of competing banks and the resulting change in the degree of bank competition mean for the firms? The effect of changing the number of banks is summarized in the following Proposition.[10]

Proposition 2. *The greater the number of banks that compete in the banking sector, the lower is the repayment and the amount of collateral, i.e.* $\frac{\partial R}{\partial N} < 0$, $\frac{\partial L}{\partial N} < 0$.

Proof: See Appendix.

Technically, a higher number of banks means that the distance which the firms have to travel to their bank decreases. Consequently, their assets have a higher liquidation value and the banks need less collateral to fulfill the break-even condition for the marginal borrower. A lower collateral requirement also reduces the repayment and increases the net return for the investing firm. However, studying the effect of market entry on bank profit, we discover a new type of externality of the entry decision of an individual bank. This is expounded in the following Proposition.

Proposition 3. *Market entry by one bank has not only a negative but also a positive externality on the return of all other banks. A positive externality arise because the average liquidation value increases as the payoff obtained from liquidating the collateralized assets of the marginal debtor increases, i.e.* $\frac{\partial \alpha^K}{\partial N} > 0$.

Proof: See Appendix.

In our model, market entry by one bank has three different effects on all other banks. First, the market share of each bank decreases. The effect on the profit of each individual bank is negative, i.e. $-\frac{1}{N^2}\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) \frac{(1-p_H)(1-\underline{\alpha})}{2(N-(1-p_H)(1-\underline{\alpha}))} < 0$. Second, the collateral requirement and the repayment decrease. With a higher number of banks, the distance between the marginal firm and the bank decreases. Therefore, the liquidation value of the assets from the marginal firm increases; and thus, the banks need less collateral in order to make zero profit. The effect of the lower collateral requirement and the lower repayment on the profit of each bank is negative, i.e. $-\frac{1}{N^2}\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) \frac{(1-p_H)(1-\underline{\alpha})(2N-(1-p_H)(1-\underline{\alpha}))}{2(N-(1-p_H)(1-\underline{\alpha}))^2} < 0$. These two effects are negative externalities and are also found in a standard Salop-model. Third, in our model there is also a positive externality. Market entry by one bank increases the average liquidation value for all banks because the distance between the bank and the firm decreases. Due to this effect, the profit of each bank increases, i.e. $\frac{1}{N^2}\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) \frac{(1-p_H)(1-\underline{\alpha})}{2(N-(1-p_H)(1-\underline{\alpha}))} > 0$. Altogether, the negative externalities outweigh the positive externality. Consequently, market entry by one bank decreases the profit of each incumbent bank.[11]

[10] Using comparative statics, we can compare different banking sectors without explicitly studying the bank's entry and exit decision.

[11] This result is not due to the specific functional form of our model. It would also be obtained if the liquidation payoff of each unit of collateral was, for instance, α^2.

Although the results in Proposition 3 are independent of the functional form chosen, it has one particularity. The negative effect of the decreasing market share on profit is just offset by the positive effect of the higher liquidation value. This is due to the fact that the change in the liquidation value is equal to the change in distance between bank and firm, which we use to measure the market share. However, it can easily be shown that the negative effects still dominate if the link between the change in market share and change in liquidation values remains linear but differs from one (as chosen in our model). To observe this, the effect that market entry has on profits by increasing the liquidation value and by decreasing the collateral requirement, as well as the repayment, are compared. The net effect of these two opposing effects is negative.

In a dynamic world the number of banks is not fixed. Potential bankers consider whether it is profitable to enter the banking sector. To establish a bank, the investor has to incur fixed costs F that include, for example, the expenditure for equipment such as computers. An investor decides to enter if

$$\Pi_i^B = \frac{1}{N}\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) \frac{(1-p_H)\left(1-\alpha^K\right)}{2\left(p_H + (1-p_H)\,\alpha^K\right)} \geq F \tag{3.1}$$

The institutional environment in transition economies is still imperfect. In the context of this model this plays a role as the costs of contract enforcement and the imperfect secondary markets influence the liquidation value of collateral α and therefore the bank's profit Π_i^B. In our model, this can be incorporated by a parameter t for the quality of the institutional environment, with $\frac{\partial \alpha}{\partial t} > 0$.

Comparative statics provide insight into the optimal number of banks (ignoring integer problems) in different institutional environments as the following Proposition shows:

Proposition 4. *In a more developed institutional environment, the equilibrium number of competing banks is lower, i.e.* $\frac{\partial N}{\partial t} < 0$.

Proof: See Appendix.

A better institutional environment means that the average liquidation value increases. This, in turn, increases bank profit. The effect of a higher average liquidation value was discusses as the third effect above. Previously, the effect was due to market entry. Qualitatively, the same result is obtained if institutions improves as $\alpha^K = \left(1 - \frac{(1-\alpha)}{N}\right)$. However, the higher liquidation value implies that the collateral requirement decreases and consequently, the repayment is lower as well - both effects decrease profit. The net effect on bank profit is negative because the impact of the lower collateral requirement/ repayment dominates as argued above. Thus, the expected profit of an entering bank is lower. Market entry is less attractive and the number of banks that enter in equilibrium is lower.

This result has interesting implications for the differences in the banking sectors in transition economies. First of all, we can explain why the number of banks varies strongly among different countries. Our model predicts that in countries with a poor institutional environment, the equilibrium number of banks is higher than in countries where the institutional environment is more developed. Second, the equilibrium number of banks should decrease in the process of transition as the quality of the institutional environment improves.

3.2. Optimal Number of Banks

The analysis so far has explained the bank's incentive to enter the market. Market entry of an additional bank has negative effects on the profit of all other banks because their market share decreases as well as their market power, which allows them to extract rents. On the other hand, market entry by one bank has a positive externality on all other banks because the average liquidation value of collateral increases. To provide an answer to the question whether transition countries are overbanked we have to study the socially optimal number of banks.

Suppose that a benevolent social planner determines the optimal number of banks N^* by maximizing the following welfare function, consisting of the net return of investment less the costs of the banking system:

$$\max_{N^*} SW = p_H X - I - E - \underbrace{\left(\frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)}{\left(p_H + (1 - p_H)\,\alpha^K\right)} (1 - p_H) \left(1 - \left(\frac{1 + \alpha^K}{2}\right)\right) + NF \right)}_{\text{costs of the banking system}} \tag{3.2}$$

Maximizing social welfare is equivalent to minimizing the costs of the banking system, which consists of the loss caused by the liquidation of collateral and the fixed costs for market entry. The following Proposition compares the number of banks that enter in equilibrium with the socially optimal number of banks:

Proposition 5. *Transition countries are underbanked, i.e.* $N^* > N$.

Proof: The following first order condition determines the optimal number of banks:

$$\frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) (1 - p_H)(1 - \underline{\alpha})}{2\left(N - (1 - p_H)(1 - \underline{\alpha})\right)^2} = F \tag{3.3}$$

Thus, the marginal increase in social welfare due to the lower loss associated with collateralization is equal to the fixed costs of market entry.

Market entry will occur until the marginal bank faces the following condition:

$$\frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right) (1 - p_H)(1 - \underline{\alpha})}{2N\left(N - (1 - p_H)(1 - \underline{\alpha})\right)} = F \tag{3.4}$$

Accordingly, banks enter until the average expected profit is equal to the fixed costs of market entry.

For a given number of banks the difference between equation (3.3) and (3.4) is given by:

$$\frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)(1-p_H)(1-\underline{\alpha})}{2\left(N-(1-p_H)(1-\underline{\alpha})\right)^2} - \frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)(1-p_H)(1-\underline{\alpha})}{2N\left(N-(1-p_H)(1-\underline{\alpha})\right)}$$

$$= \frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)(1-p_H)(1-\underline{\alpha})}{2N\left(N-(1-p_H)(1-\underline{\alpha})\right)^2} > 0$$

As the effect of market entry on social welfare is higher than on the individual bank's profit and as both functions are concave, the socially optimal number of banks, N^*, is higher than the number of banks entering in equilibrium, N. Therefore, transition countries are underbanked. Q.E.D.

There are two reasons why the individual decision for market entry diverges from the social choice. First, the bank's decision about market entry is orientated towards the average profit to be expected. In contrast, the social choice is determined by the marginal effect of market entry on social welfare. Second, the social welfare function differs from the profit function. In our model, social welfare is reduced by the costs of collateralization. These costs decrease when the average loss of liquidation decreases and when the amount of collateral decreases. As market entry reduces the collateral requirement and the average loss of liquidation, entry influences social welfare positively through both "channels". And these positive influences on social welfare are the reason why a social planner wants a higher number of banks. Compared to the social planner's decision, the individual bank neglects the positive effects of its decision on social welfare. First, market entry by one bank increases the liquidation value for all banks, and thereby social welfare. However, the entrant bank considers only the positive effect on its own profit. Second, market entry reduces the amount of collateral, which reduces the costs of collateralization and thereby increases social welfare. However, for the entrant bank less collateral decreases its expected profit. As a consequence, entry remains insufficient.[12]

4. Discussion and Conclusion

The prediction of our model is that the number of banks decreases as the institutional environment improves. What has been observed in transition countries is that the number of banks soared in the first years of transition but slightly decreased later (EBRD, 2002). The decrease observed recently might be explained by the better institutions that limit the bank's scope for extracting rents.

The other important result of our model is that transition countries are underbanked. The evidence from Jaffee and Levonian (2001) is that transition countries can be overbanked as well as underbanked (see Table 1 in the Appendix). According to their calculation the Czech Republic, Hungary and Poland are underbanked. Whereas all other

[12] Matsumura (2000) shows that the excess-entry theorem does not always hold if the integer problem is considered and marginal production cost is increasing. He warns against using the excess entry theorem precipitately, i.e. its application in the Japanese Large-Scale Retail Act which restricts the new entry of retailers.

countries, among them Russia, are overbanked. How can we explain the difference between the result of Jaffee and Levonian and our theoretical prediction? First of all, their benchmark calculation depends only on the level of GDP, neglecting other important factors such as the institutional environment. The socially optimal number of banks increases as institutions deteriorate. Thus, their benchmark should be lower than the socially optimal number of banks of our analysis. Second, the fixed costs that banks face might be lower than the actual fixed costs of establishing a bank. In Russia, for instance, many banks are founded by firms. These so-called pocket banks get access to premises or office facilities more easily and more cheaply because they are leased below market prices. Therefore, their costs of market entry are lower than the actual costs, which are considered by a social planner. Under such circumstances, the incentive for market entry increases. The gap between the equilibrium and the optimal number of banks decreases as the institutions improve.

What are the effects of underbanking? We have argued that a lower number of banks increases the collateral requirement. We did not explicitly model credit rationing. However, it is obvious that some firms would lose access to financing if the collateral requirement becomes more demanding. Moreover, we should take into account that information about the asset endowment is asymmetric. A bank's reaction to the lack of information could be credit rationing which increases with a higher collateral requirement (Hainz, 2003). The negative consequences of underbanking can be reduced if the institutional environment improves. Therefore, the following measures should be considered. First, the laws on collateralization must be drafted carefully and unambiguously. Second, public contract enforcement has to become cheaper and more reliable. This would reduce the incentive for private contract enforcement; under which a bank can extract rents since it possesses a better contract enforcement technology than its competitors. Finally, the development of secondary markets has to be fostered. Improving the legal framework not only reduces the costs of collateralization by facilitating its seizure but also by increasing the efficiency of secondary markets.

5. Appendix

5.1. Proof of Proposition 1

(1) The firm
In our model the banks cannot price-discriminate. Therefore, we have to find the contract which specifies the most favorable contract terms $\{R, L\}$ for the marginal firm.

(2) The bank i
We first specify the contract for the marginal firm that yields zero profit for the bank. We then show that it is optimal to offer this contract.

As we have assumed that the banks are equidistantly distributed, and α_i decreases proportionally with the distance travelled, the marginal firm where $\alpha_i = \alpha_j \left(= \alpha^K\right)$ is located midway between the two banks. Due to (IC-F) and $\alpha^K \leq 1$, bank i sets $R = X + L - \frac{E}{\Delta p}$. From the bank's zero profit constraint of the marginal firm:

$$\Pi_i^B = p_H \left(X + L - \frac{E}{\Delta p}\right) + (1 - p_H)\, \alpha^K L - I = 0, \tag{5.1}$$

the optimal collateral is obtained $L = \frac{\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)}{p_H + (1-p_H)\alpha^K}$. According to (IC-F) $R = \frac{\left(X - \frac{E}{\Delta p}\right)(1-p_H)\alpha^K + I}{p_H + (1-p_H)\alpha^K}$. Due the assumptions on the size of A, (PC-F) is fulfilled.

If both banks offer the same credit contract as specified above, they have the same market share, the profit from the marginal firm is zero and the total profit of each bank is positive. Deviating from this contract reduces a bank's profit. On the one hand, bank i does not have an incentive to increase the collateral requirement and the repayment because it would lose all its customers and make zero profit. On the other hand, bank i does not have an incentive to decrease the collateral requirement and the repayment because it would attract firms that are further away than the marginal firm. Profit generated from the marginal firm profit is zero. Thus, the profit generated from firms that are located further away would be negative. Q.E.D.

5.2. Proof of Proposition 2

$\frac{\partial L}{\partial N} = -\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)(1 - p_H)\, \frac{1-\underline{\alpha}}{(N-(1-p_H)(1-\underline{\alpha}))^2} < 0$
$\frac{\partial R}{\partial N} = -\left(-p_H X + I + \frac{E}{\Delta p} p_H\right)(1 - p_H)\, \frac{1-\underline{\alpha}}{\Delta p(N-(1-p_H)(1-\underline{\alpha}))^2} < 0$ Q.E.D.

5.3. Proof of Proposition 3

$\frac{d\Pi_i^B}{dN} = -\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)(1-p_H)(1-\underline{\alpha})\frac{2N-(1-p_H)(1-\underline{\alpha})}{2N^2(N-(1-p_H)(1-\underline{\alpha}))^2} < 0$

To determine the sign of $2N - (1-\underline{\alpha})(1-p_H)$ we have to determine the equilibrium number of banks. Bank decide to enter until

$\left(\frac{\left(-p_H X+I+\frac{Ep_H}{\Delta p}\right)}{N}\right)\frac{(1-p_H)\left(1-\left(1-\frac{(1-\underline{\alpha})}{N}\right)\right)}{2\left(p_H+(1-p_H)\left(1-\frac{(1-\underline{\alpha})}{N}\right)\right)} - F = 0.$

Thus, the equilibrium number of banks in the market is:

$$\begin{aligned} N_1 &= \frac{1}{2F}(1-p_H)(1-\underline{\alpha})F + \\ &\quad \frac{1}{2F}\sqrt{(1-p_H)(1-\underline{\alpha})F\left((1-p_H)(1-\underline{\alpha})F + 2\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\right)} \\ N_2 &= \frac{1}{2F}(1-p_H)(1-\underline{\alpha})F - \\ &\quad \frac{1}{2F}\sqrt{(1-p_H)(1-\underline{\alpha})F\left((1-p_H)(1-\underline{\alpha})F + 2\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\right)} \end{aligned}$$

where N_2 can be excluded because it would be negative.

Therefore, we conclude that $2N - (1-\underline{\alpha})(1-p_H) > 0$.

The total effect of market entry on bank profit consists of the following effects:

(1) increasing liquidation value α^K (for given market share, repayment and collateral):

$\left(\frac{1}{N^2}\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\frac{(1-p_H)(1-\underline{\alpha})}{2(N-(1-p_H)(1-\underline{\alpha}))}\right) > 0$

(2) decreasing collateral L and repayment R (for given market share and liquidation value):

$\left(-\frac{1}{N^2}\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\frac{(1-p_H)(1-\underline{\alpha})(2N-(1-p_H)(1-\underline{\alpha}))}{2(N-(1-p_H)(1-\underline{\alpha}))^2}\right) < 0$

(3) decreasing market share (for given liquidation value, repayment and collateral):

$\left(-\frac{1}{N^2}\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\frac{(1-p_H)(1-\underline{\alpha})}{2(N-(1-p_H)(1-\underline{\alpha}))}\right) < 0$ Q.E.D.

5.4. Proof of Proposition 4

Totally differentiating Π_i^B we obtain

$\frac{dN}{dt} = -\frac{\frac{d\Pi_i^B}{dt}}{\frac{d\Pi_i^B}{dN}} < 0$ as

$\frac{d\Pi_i^B}{dt} = -\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)\frac{(1-p_H)\frac{\partial\underline{\alpha}}{\partial t}}{2(N-(1-p_H)(1-\underline{\alpha}))^2} < 0$

$\frac{d\Pi_i^B}{dN} = -\left(-p_H X + I + \frac{Ep_H}{\Delta p}\right)(1-p_H)(1-\underline{\alpha})\frac{2N-(1-p_H)(1-\underline{\alpha})}{2N^{2(N-(1-p_H)(1-\underline{\alpha}))2}} < 0$ Q.E.D.

5.5. Table 1

Table 1: Analyzing The Number of Banks in Transition Economies

	Actual Number	Benchmark Number	GDP per Capita	EBRD Legal Transition Indicator
Poland	81	82	3487	4
Czech Republic	51	55	4885	4
Slovak Republic	29	25	3443	3
Slovenia	29	22	10499	3
Lithuania	12	7	2434	3
Hungary	43	62	6019	4
Belarus	38	21	2329	2
Ukraine	188	101	3042	2
Uzbekistan	29	13	1014	2
Bulgaria	47	17	2305	3
Estonia	14	5	3055	2
Albania	9	3	1751	2
Kazakstan	101	29	1963	2
Mongolia	13	3	911	2
Krygz Republic	18	4	1746	2
Latvia	33	7	3707	3
Russia	2030	367	3983	3
Croatia	60	9	4266	4
Armenia	33	3	1425	3
Georgia	61	3	1151	2
Azerbaijan	136	5	1174	1

Source: Jaffee and Levonian (2001), p. 169.

6. References

Bester, Helmut, "Screening and Rationing in Credit Markets with Imperfect Information." *American Economic Review* **75**, 4: 850 - 855, September 1985.

Bester, Helmut, "The Role of Collateral in a Moral Hazard Model of Debt Renegotiation." *Journal of Money, Credit, and Banking* **26**, 1: 72 - 86, February 1994.

Bonin, John P., Miszei, Kálmán, Székely, István P., and Wachtel, Paul, *Banking in Transition Economies: Developing Market Oriented Banking Sectors in Eastern Europe.* Cheltenham and Northampton: Edward Elgar, 1998.

Bonin, John and Wachtel, Paul, "Toward Market-Oriented Banking in the Economies of Transition." In Mario J. Brejer and Marko Skreb, Eds., *Financial Sector Transformation, Lessons from Economies in Transition*, pp. 93 - 131. Cambridge, UK: Cambridge University Press, 1999.

Bratkowski, Andrzej, Grosfeld, Irena, and Rostowski, Jacek, "Investment and Finance in De Novo Private Firms: Empirical Results from the Czech Republic, Hungary and Poland." *Economics of Transition* **8**, 1: 101 - 116, 2000.

Broecker, Thorsten, "Credit-Worthiness Tests and Interbank Competition." *Econometrica* **58, 2:** 429 - 452, March 1990.

Degryse, Hans and Ongena, Steven, "Distance, Lending Relationships, and Competition," mimeo, Katholieke Universiteit Leuven, March 2002.

Dittus, Peter and Prowse, Stephen, "Corporate Control in Central Europe and Russia. Should Banks Own Shares?" In Roman Frydman, Cheryl W. Gray, and Andrzej Rapaczynski, Eds., *Corporate Governance in Central and Eastern Europe and Russia.* Vol. 2., Insiders and the State, pp. 20 - 67. Budapest et al.: Central European University Press, 1996.

EBRD, *Transition Report 2000: Employment, Skills and Transition*, EBRD: London, 2000.

EBRD, *Transition Report 2002: Agriculture and Rural Transition*, EBRD: London, 2002.

Fan, Qimiao, Lee, Une F., and Schaffer, Mark E., "Firms, Banks, and Credit in Russia." In Simon Commander, Qimiao Fan, and Mark E. Schaffer, Eds., *Enterprise Restructuring and Economic Policy in Russia,* pp. 140 - 165. Washington D.C.: EDI Development Studies, 1996.

Gabszewicz, Jean Jaskold and Thisse, Jacques-Francois, "Spatial Competition and the Location of Firms." In Arnott Richard, Ed., *Location Theory. Fundamentals of Pure and Applied Economics*, pp. 1 - 71, London: Harwood Academic Press, 1986.

Gromberg, Timothy and Meyer, Jack, "Competitive Equilibria in Uniform Delivered Pricing Models." *American Economic Review* **71**, 4: 758 - 763, 1981.

Hainz, Christa, "Bank Competition and Credit Markets in Transition Economies", Journal of Comparative Economics, forthcoming 2003.

Holmström, Bengt, "Financing of Investment in Eastern Europe: A Theoretical Perspective." *Industrial and Corporate Change* **5**, 2: 205 - 237, 1996.

Jaffee, Dwight and Levonian, Mark, "Structure of Banking Systems in Developed and Transition Economies", *European Financial Management* **7**, 2: 161 - 181, June 2001.

Kahn, Peter L., "The Russian Bailiffs Service and the Enforcement of Civil Judgments." *Post-Soviet Affairs* **18**, 2: 148 - 181, 2002.

Koford, Kenneth and Tschoegl, Adrian E., "Problems of Bank Lending in Bulgaria: Information Asymmetry and Institutional Learning." *Moct-Most* **9**, 123 - 151, 1999.

Matsumura, Toshihiro, "Entry Regulation and Social Welfare with an Integer Problem." *Journal of Economics* **71**, 1: 47 - 58, 2000.

McMillan, John and Woodruff, Christopher, "Privat Order under Dysfunctional Public Order." *Michigan Law Review* **98**, 8: 2421 - 2458, August 2000.

McNulty, James E. and Harper, Joel T. , "Legal Systems, Financial Intermediation and the Development of Loan Relationships in the Transitional Economies of Central and Eastern Europe," mimeo, Florida Atlantic University, December 2001.

OECD, *Economic Surveys: Russian Federation.* Paris: OECD, 2000.

Salop, Steven, "Monopolistic Competition with Outside Goods." *Bell Journal of Economics* **10**, 141 - 156, 1979.

Schnitzer, Monika, "Bank Competition and Enterprise Restructuring in Transition Economies." *Economics of Transition* **7**, 1: 133 - 155, 1999.

General Equilibrium Model Of An Economy With A Futures Market

Roman Cech
Osteuropa-Institut München

December, 2002

Abstract

General equilibrium effects of a futures market are analyzed in a two-sector model of an economy populated with agents who have differential risk aversion. Conditions leading to changes in the industry formation are derived and their effect on agents' welfare is measured by equivalent variation. A class of speculators who most benefit from the futures market endogenously arises in equilibrium.

1 Introduction

This paper presents a general equilibrium model of an economy with futures markets, in which futures trading arises among agents whose only heterogeneity is their degree of risk aversion. The economy produces two goods. The production is deterministic in one industry, but it is random in the other. Agents specialize in the production of one of the two goods. While all agents are exposed to the same price risk as consumers of the two goods, they are exposed to a different, industry-specific, income risk as producers. They choose which industry to enter based on the income risk. When given the option of sharing risk through the futures market, agents trade futures for two reasons: because they have a differential risk aversion, and because they have differential income risk exposure.

The General Equilibrium pricing of derivative securities, including futures contracts, is well understood in two cases: when there exists a complete set of contingent claims markets (Arrow, 1953), or when agents are identical (Lucas, 1978). Unfortunately, these models are too simple to study many aspects of futures trading, because real world asset markets are incomplete and agents are not identical. By contrast, in the model presented here, agents differ in their risk aversion and risk exposure, and asset markets are incomplete.

In spirit, this model is an extension of Kihlstrom and Laffont (1979). In that model, there exists a continuum of agents differing only in their degree of risk aversion who decide whether to become entrepreneurs or workers. Our model, like that of Kihlstrom and Laffont, uses a continuum of agents with differential risk aversion to explore industry choice and futures trading.

It seems natural to link the agents' industry choice and their behavior in the futures market to their risk aversion. Intuitively, risk sharing is what futures trading is all about. Yet, most of the literature on the topic is based on partial equilibrium analysis, with little, if any, role for risk sharing.[1] The few general equilibrium models of futures markets do not shed more light on the crucial role of risk sharing than their partial equilibrium counterparts. They generate different patterns of behavior by assigning speculators, hedgers, producers, etc. entirely different objective functions and different access to information. Speculation, hedging, industry choice and futures trading all arise endogenously in this model, whereas they are artificially imposed in the existing literature.

The modern thought on backwardation and contango[2] dates back to J. M. Keynes (1930, p.144), who wrote: "The quoted forward price,..., must fall below the anticipated future spot price by at least the amount of the normal backwardation." He considered backwardation an equivalent of a risk premium. Assuming that most hedgers were taking short positions, he concluded that the futures price has to be below the expected spot price in order to attract a sufficient number of long speculators to clear the market.

Formal models have been developed in an attempt to confirm or refute Keynes's conjecture. Two previous general equilibrium models have shown, as this paper does, that either backwardation or contango may occur. The previous models, however, are rather cumbersome with a very complicated structure. Danthine's (1978) seminal paper also uses agents who are exogenously destined to be either speculators or hedgers. Danthine shows that the bias is generated when speculators are endowed with more information than hedgers. The differential is a source of the speculators' expected profit. In our model, there is no exogenous industry formation or exogenous division of population. On the contrary, speculation arises endogenously.

Anderson and Danthine (1983) showed conditions that may lead to backwardation and contango. Results they obtained reflect a complicated exogenously-determined market structure. The Anderson and Danthine model is populated by four types of agents: speculators, producers, processors and storage companies. Each type has a different objective function. Uncertainty enters their model via two exogenous production and demand shocks. In contrast, the model presented in this paper generates both kinds of bias in an economy with agents who have identical objective functions, except for their degree of risk aversion. Due to the general equilibrium nature of our model, agents are exposed to the price risk and income risk generated by a single source of uncertainty, a production shock in one of the two industries.

Conventionally, commodity producers are expected to hedge by selling their output at a predetermined price. If they do not hedge or if they speculate by taking long positions, it is taken as evidence that they are risk loving, or they have access to private information.

[1] Optimal hedging and speculation in the presence of joint income and output risk have been studied by Grant (1985). Other related papers include Briys and Schlesinger (1993) who use state dependent preferences, Karp (1987) who studies dynamic hedging with stochastic production, Lapan, Moschini and Hanson (1991) consider simultaneous presence of futures and options markets. Haruna (1996) studies relationship between spot and futures prices in a simple long-run competitive industry partial equilibrium context with no exogenous speculators.

[2] Backwardation (contango) occurs when the current cash price is greater than (is less than) the futures price. In this model, backwardation (contango) refers to the relationship between the futures price and the expected spot price.

Results of this paper contradict conventional wisdom in two respects. First, what appears to be speculation may in fact be the behavior of a very risk averse individual. In our model, with all agents having the same information, a surprising and seemingly counter-intuitive pattern of futures trading arises as a result of combined exposure to price and income risk. The oversimplified conventional view of hedging and speculation has been pointed out by Fouda at al. (1999), who showed that different producers may take either short or long futures positions because they are exposed to a differential cash flow risk. In their model, producers differ from each other in the technology they use to produce the commodity. Risk aversion, however, plays no role in their results. In this paper, the risk exposure of all commodity producers is identical; they choose different futures positions because of their differential risk aversion.

Second, in general, there are agents who are worse off in terms of their expected utility when the futures market is available. The possibility of risk sharing in the futures market attracts entry of agents who would otherwise find the industry with stochastic output too risky. That leads to a greater output and a lower relative price of that good. As a result, other producers of that good who choose not to trade when futures trading is available must be worse off, because they sell their product at a lower spot price.

The existing theory of futures markets pays almost no attention to their welfare effects. The existence of futures markets is usually treated as ex-ante beneficial to all agents. Although real-world commodity producers have sometimes complained that futures markets harm them, the only rationale for this in existing models is ex-post price regret. In partial equilibrium, welfare does not seem to be an interesting issue because every producer can choose to stay away from the futures market and be no worse off than he would be in the absence of it. This view, however, neglects the effect futures markets may have on the long-run industry formation and subsequently on relative prices. If industry formation changes, the resulting change in relative prices will impact the welfare of virtually everyone in the economy. In this model, equivalent variation is used to study the effect of the futures market on the welfare of individual agents. We show that there may exist an identifiable group of endogenous speculators, and that when such a group arises, they benefit most.

The rest of the paper is organized in three sections. Equilibrium properties of the model with the aggregate production shock are shown in Section 2. The effect of the idiosyncratic production shock is analyzed in Section 3. The welfare effect of the futures market is investigated in Section 4.

2 Model With An Aggregate Production Shock

In this section, we study properties of the economy in which agents make all their production and consumption decisions in one period. Strong closed-form results are the benefit of the simplicity of the model.

Population

The economy is populated by a continuum of agents who have identical preferences except for their coefficient of absolute risk aversion. The agent's risk aversion is determined by his type, $a \in (\mu_a - \lambda_a, \mu_a + \lambda_a)$, where $0 < \mu_a - \lambda_a < \mu_a + \lambda_a < \infty$. Agents are uniformly distributed on the interval $(\mu_a - \lambda_a, \mu_a + \lambda_a)$; μ_a is the average degree of risk aversion and λ_a is the dispersion of risk aversion. The size of the population is normalized to 1.

Preferences

There are two goods, x and y. The utility function of an agent of type a is[3]

$$u = -\exp\left[-ac_x^{\beta}c_y^{1-\beta}\right],$$

where c_x is the consumption of good x and c_y is the consumption of good y.

Parameter β is the same for all agents. It determines the proportion of income agents spend on good x. Good x is a numeraire, its price is normalized to 1. Demand functions for goods x and y are $c_x = \beta I$, $c_y = \frac{(1-\beta)I}{p}$, where I is the agent's income. The relative price of good y, p, is expressed in units of good x. Substituting demand functions into the utility function gives the indirect utility function

$$V = \exp\left[-a\frac{\widehat{\beta}}{p^{1-\beta}}I\right],$$

where $\widehat{\beta} \equiv \beta^{\beta}(1-\beta)^{1-\beta}$.

Production And Income

Each agent possesses an indivisible unit of labor and supplies it inelastically. The agent must specialize in the production of either x or y. In the x-industry, the unit of labor is transformed into 1 unit of good x (a harmless normalization). If the agent devotes his unit of labor to the production of y, output, $\widetilde{Y} = \mu + \tilde{\lambda}$, is a random variable. For the sake of simplicity, it is assumed that $\widetilde{Y}$ has a Bernoulli distribution with the high state (good harvest) $\widetilde{Y} = \mu + \lambda$ occurring with probability 1/2 and the low state (bad harvest) $\widetilde{Y} = \mu - \lambda$ occurring with probability 1/2, where $\mu - \lambda > 0$.[4] The random shock is an aggregate shock; all of the y-producers have a good harvest or all of them have a bad one. Each agent sells the output at the market price and generates income $I_x = 1$, or $I_y = \widetilde{p}\widetilde{Y}$, respectively.

2.1 Industry Formation Without A Futures Market

Rational Expectations

Each agent chooses to enter one of the two industries. Then he produces in the industry of his choice, observes the realization of the random shock, sells his output in the spot market and consumes his preferred consumption bundle.

[3] The validity of this approach to represent risk aversion of agents consuming many commodities has been established in Kihlstrom and Mirman (1974).

[4] Since a more general distribution of the two production states with probabilities q and $(1-q)$ does not affect the qualitative nature of results of this model, $q = 1/2$ is used for simplicity throughout this dissertation.

At the time agents choose the industry they will enter, they do not know the realization of $\widetilde{Y}$ or $\widetilde{p}$. Agents therefore choose the industry which promises greater expected utility EV:

$$EV^x = -\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\right]-\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\right], \quad (1.a)$$

$$EV^y = -\frac{1}{2}\exp\left[-a\widehat{\beta}\underline{p}^{\beta}(\mu+\lambda)\right]-\frac{1}{2}\exp\left[-a\widehat{\beta}\overline{p}^{\beta}(\mu-\lambda)\right] \quad (1.b)$$

The low price $\underline{p}$ corresponds to the high output state and the high price $\overline{p}$ to the low output state of the world. Agents' expectations are rational: the resulting distribution of relative price is as they expected.

Market Equilibrium

Let A be the measure of the subset of agents who produce x; $(1-A)$ is a measure of the subset of agents who produce y. Quantity demanded of good x is the sum of quantities demanded by x-producers and y-producers: $\beta A+\beta(1-A)\widetilde{p}\widetilde{Y}$. The supply of good x is equal to A. The market for good x clears when $\beta A+\beta(1-A)\widetilde{p}\widetilde{Y}=A$. This reduces to $\widetilde{p}=\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\frac{1}{\widetilde{Y}}$. When the market for x clears, the market for good y must clear by Walras law. The distribution of relative price $\widetilde{p}$ of good y is then

$$\widetilde{p}=\begin{cases}\underline{p}=\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\frac{1}{(\mu+\lambda)}, \text{with probability } \frac{1}{2}\\ \overline{p}=\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\frac{1}{(\mu-\lambda)}, \text{with probability } \frac{1}{2}\end{cases}. \quad (2)$$

Substituting expressions (2) into the utility function EV^y in (1.b) gives

$$EV^y=-\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\right]-\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\right]. \quad (1.c)$$

Definition 1 *The population is divided into a subset G of x-producers and a subset H of y-producers. An equilibrium is a set $(\{G,H\},\widetilde{p})$, where $\{G,H\}$ is a partition of the population such that $EV^x\geq EV^y$ holds for all $a\in G$, and $EV^x\leq EV^y$ holds for all $a\in H$. Relative price $\widetilde{p}$ satisfies (2), where the proportion of x-producers $A=\int_G d\nu(a)$ and the proportion of y-producers $(1-A)=\int_H d\nu(a)$. Measure ν is a uniform measure on $(\mu_a-\lambda_a,\mu_a+\lambda_a)$.*

Proposition 2 *An equilibrium partition $\{G,H\}$ is any partition that satisfies $A=\beta$. The equilibrium price in the two states of the world is $\overline{p}=\frac{1}{(\mu-\lambda)}$ and $\underline{p}=\frac{1}{(\mu+\lambda)}$, respectively.*

Proof. Note that the two expected utility functions (1.a) and (1.c) are equal to each other (for all a) if and only if $1=\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}$. This implies that $A=\beta$. If $A<(>)\beta$, all agents would want to enter the x-industry (y-industry), which is not consistent with equilibrium. The corresponding equilibrium price in the two output states is obtained by substituting $A=\beta$ in (2). This completes the proof. ■

If $A = \beta$, all agents are indifferent between entering the two industries. Any partition of agents that results in β agents producing x and $(1-\beta)$ agents producing y is consistent with equilibrium. This strong result is due to the unitary price elasticity of demand that eliminates income uncertainty and due to the fact that the production shock is the same for all producers. Recall that the x-producer's income is $I^x = 1$. In equilibrium, the y-producer's income is $I^y = \underline{p}\,(\mu + \lambda) = 1$ when the output is high and $I^y = \overline{p}\,(\mu - \lambda) = 1$ when the output is low. The fact that it is the y-industry that has the random output does not make it a "riskier" industry. As producers, all agents have the same income. As consumers, they face the same price uncertainty.

Note that the equilibrium proportion of x-producers and y-producers in the population is determined by the split of agents' income. Portion βI is used to buy x and portion $(1-\beta)\,I$ is used to buy y. It is not affected by the presence of the aggregate shock; it would be the same if output were deterministic in both industries .

2.2 Industry Formation With A Futures Market

Suppose that, at the time he chooses the industry, each agent has the opportunity to trade futures on commodity y. Since the only source of risk here is the aggregate shock resulting in two states of the world, the futures contract is equivalent to a contingent claims contract.

Let z be the number of units of y an agent contracts to buy and let p_F be the futures price. Agents who choose $z > 0$ take long positions (they will buy z units of y at the price p_F), while those who choose $z < 0$ take short positions (they will sell $|z|$ units of y at the price p_F).[5] The delivery will take place after realizations of the random output $\widetilde{Y}$ and random price $\widetilde{p}$ have been observed. The income of an agent is then given by

$$I^x = 1 + z(\widetilde{p} - p_F), \text{ or} \tag{3.a}$$

$$I^y = pY + z(\widetilde{p} - p_F), \tag{3.b}$$

respectively.

Choice Of The Futures Position z

Let $\overline{D} \equiv \overline{p} - p_F$ and $\underline{D} \equiv p_F - \underline{p}$. The expected utility function of an agent who enters the x-industry is

$$EV^x = \underset{z}{MAX}\left\{-\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\left(1 - z\underline{D}\right)\right] - \frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\left(1 + z\overline{D}\right)\right]\right\}, \tag{4.a}$$

while the y-producer's utility function is

$$EV^y = \underset{z}{MAX}\left\{\begin{array}{c} -\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\left(\underline{p}\,(\mu+\lambda) - z\underline{D}\right)\right] \\ -\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\left(\overline{p}\,(\mu-\lambda) + z\overline{D}\right)\right] \end{array}\right\}. \tag{4.b}$$

[5] The contracts are referred to as futures even though the setup is too general to distinguish them from forward contracts.

Both functions are concave in z. Agents choose their optimal futures positions

$$z^x = \arg\max EV^x \text{ and } z^y = \arg\max EV^y$$

subject to the constraint $I^x, I^y \geq 0$. Any potential loss resulting from a futures position must be backed by income from production because agents are not allowed to default on their obligation in any state of the world. When the solutions are interior, first-order conditions give

$$z^x = \frac{\left(\frac{\widehat{\beta}}{\underline{p}^{1-\beta}} - \frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\right) - \frac{1}{a}\ln K}{\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\overline{D} + \frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\underline{D}}, \text{ and} \tag{5.a}$$

$$z^y = \frac{\left(\widehat{\beta}\underline{p}^{\beta}(\mu+\lambda) - \widehat{\beta}\overline{p}^{\beta}(\mu-\lambda)\right) - \frac{1}{a}\ln K}{\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\overline{D} + \frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\underline{D}}, \tag{5.b}$$

where $K \equiv \frac{\underline{D}}{\overline{D}}\left(\frac{\overline{p}}{\underline{p}}\right)^{1-\beta} \equiv \frac{\left(p_F - \underline{p}\right)}{\left(\overline{p} - p_F\right)}\left(\frac{\overline{p}}{\underline{p}}\right)^{1-\beta}$. Substituting optimal futures positions (5.a) and (5.b) into the functions (4.a) and (4.b) gives

$$EV^x = -2MK^{1+M}\exp\left[-a\frac{\widehat{\beta}\left(\overline{p} - \underline{p}\right)}{\overline{D}\underline{p}^{1-\beta} + \underline{D}\overline{p}^{1-\beta}}\right], \text{ and} \tag{6.a}$$

$$EV^y = -2MK^{1+M}\left(\exp\left[-a\frac{\widehat{\beta}\left(\overline{p} - \underline{p}\right)}{\overline{D}\underline{p}^{1-\beta} + \underline{D}\overline{p}^{1-\beta}}\right]\right)^{\left(\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}\right)}, \tag{6.b}$$

where $M \equiv \frac{\frac{\overline{D}}{\overline{p}^{1-\beta}}}{\frac{\overline{D}}{\overline{p}^{1-\beta}} + \frac{\underline{D}}{\underline{p}^{1-\beta}}}$.

Notice that EV^y differs from EV^x only by the exponent of $\frac{(1-\beta)}{\beta}\frac{A}{(1-A)}$. If the exponent is greater (smaller) than one, all agents want to enter the y-industry (x-industry), which is not consistent with the equilibrium. All agents are indifferent between the two industries if and only if $\frac{(1-\beta)}{\beta}\frac{A}{(1-A)} = 1$.

Equilibrium in the Futures Market

The futures market clears when every long futures position is offset by a short futures position. The sum of all positions must therefore add up to zero. If the sum of the futures positions held by the x-producers is $\int_G z^x d\nu(a)$ and the sum of the futures positions held by the y-producers is $\int_H z^y d\nu(a)$, then the futures price p_F must solve

$$\int_G z^x d\nu(a) + \int_H z^y d\nu(a) = 0. \tag{7}$$

Equilibrium in the Spot Market

Market demand for good x consists of the quantity

$$\beta \int_G (1 + (\widetilde{p} - p_F) z^x)\, d\nu(a)$$

demanded by x-producers and the quantity

$$\beta \int_H \left(\widetilde{p}\widetilde{Y} + (\widetilde{p} - p_F) z^y\right) d\nu(a)$$

demanded by y-producers. The supply of good x is A. Quantity demanded of good x equals quantity supplied when $\widetilde{p}$ solves

$$\beta \int_G (1 + (\widetilde{p} - p_F) z^x)\, d\nu(a) + \beta \int_H (pY + (\widetilde{p} - p_F) z^y)\, d\nu(a) = A.$$

This can be written as

$$\beta A + \beta\,(1 - A)\, pY + \beta(\widetilde{p} - p_F) \left[\int_G z^x d\nu(a) + \int_H z^y d\nu(a) \right] = A.$$

When the futures market clears, (7) holds and the spot market clearing condition reduces to

$$\beta A + \beta\,(1 - A)\, pY = A.$$

The equilibrium relative price $\widetilde{p}$ then satisfies Eq. (2):

$$\widetilde{p} = \begin{cases} \underline{p} = \frac{(1-\beta)}{\beta} \frac{A}{(1-A)} \frac{1}{(\mu+\lambda)}, \text{with probability } \frac{1}{2} \\ \overline{p} = \frac{(1-\beta)}{\beta} \frac{A}{(1-A)} \frac{1}{(\mu-\lambda)}, \text{with probability } \frac{1}{2} \end{cases}.$$

Definition 3 *Equilibrium is a set* $(\{G, H\}, \widetilde{p}, p_F)$*, where* $\{G, H\}$ *is a partition of the population such that* $EV^x \geq EV^y$ *holds for all* $a \in G$*, and* $EV^x \leq EV^y$ *holds for all* $a \in H$*. Relative price* $\widetilde{p}$ *satisfies Eq. (2) and futures price* p_F *satisfies Eq. (7); both the spot market and the futures market clear.*

When the no-default constraint is not binding, equilibrium is described in the following proposition.

Proposition 4 ***(A)*** *The equilibrium partition is any* $\{G, H\}$ *such that* $A = \beta$. *Equilibrium prices in the two states of the world are* $\overline{p} = \frac{1}{(\mu-\lambda)}$ *and* $\underline{p} = \frac{1}{(\mu+\lambda)}$*, respectively.*

(B) *Each agent's choice of futures position is independent of the industry he enters. There is a unique agent* $\hat{a} = \frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}$ *who chooses* $z = 0$ *(see Figure 1). All* $a < (>)\,\hat{a}$ *choose* $z < (>)0$*; more risk-averse agents take long positions and less risk-averse agents take short positions.*

(C) *The equilibrium futures price is*

$$p_F = \left\{ \frac{1}{1+\left(\frac{\mu+\lambda}{\mu-\lambda}\right)^{1-\beta}\left(\frac{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu+\lambda)^{1-\beta}\right]}{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu-\lambda)^{1-\beta}\right]}\right)} \right\} \frac{1}{(\mu-\lambda)}$$
$$+ \left\{ \frac{\left(\frac{\mu+\lambda}{\mu-\lambda}\right)^{1-\beta}\left(\frac{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu+\lambda)^{1-\beta}\right]}{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu-\lambda)^{1-\beta}\right]}\right)}{1+\left(\frac{\mu+\lambda}{\mu-\lambda}\right)^{1-\beta}\left(\frac{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu+\lambda)^{1-\beta}\right]}{\exp\left[-\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)}\widehat{\beta}(\mu-\lambda)^{1-\beta}\right]}\right)} \right\} \frac{1}{(\mu+\lambda)}.$$

(D) *The equilibrium futures price* p_F *is an increasing function of the average coefficient of risk aversion* μ_a *and of the aggregate shock* λ. *It is a decreasing function of the mean output* μ.

(E) *The futures price is equal to the expected spot price if and only if*

$$\frac{2\lambda_a}{\ln(\mu_a+\lambda_a)-\ln(\mu_a-\lambda_a)} = \frac{\ln(\mu+\lambda)^{1-\beta}-\ln(\mu-\lambda)^{1-\beta}}{\widehat{\beta}\,(\mu+\lambda)^{1-\beta}-\widehat{\beta}\,(\mu-\lambda)^{1-\beta}}.$$

If the left-hand side is smaller (greater) than the right-hand side, then $p_F > (<)\, E(p)$.

Proof. See Appendix A. ■

Figure 1 depicts each agent's futures position as a function of his coefficient of absolute risk aversion, which is on the horizontal axis. Increasing values of a represent greater risk aversion. Positive futures positions are long, negative positions are short. The area under the curve represents the sum of all futures positions taken by agents of different types. In equilibrium, the area enclosed by the long positions has to be equal in size to the area enclosed by the short positions.

The intuition behind the equilibrium in the product market is the same as in the previous section. The result is again driven by the unitary price elasticity of demand and the aggregate nature of the production shock. Agents in both industries face the same price risk. The income from production, measured in units of good x, is certain and is the same in both industries.

The futures market gives more risk averse agents the opportunity to insure against the worst-case scenario, which is the state of the world where the output of y is low and

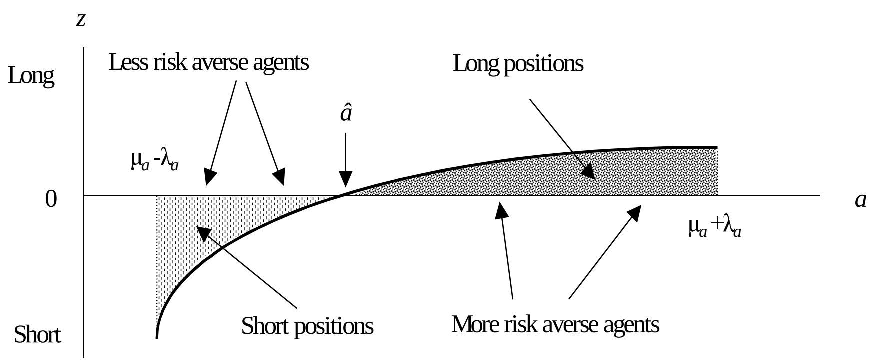

Figure 1: Futures positions held by agents

its price is high. In this case, it is better to buy y at the price $p_F < \bar{p}$. For a risk averse agent, the additional income from futures trading in this bad state compensates for the high price of y. That is why more risk-averse agents take long positions. The less risk-averse agents, on the other hand, take the opportunity to increase their expected utility by taking short positions and exposing themselves to a greater risk in utility.

Properties Of The Equilibrium Futures Price p_F

The monotonicity of p_F in μ, λ, μ_a and λ_a is intuitive. First, the relative price of y decreases with the mean output μ in every state of the world. That is why the futures price decreases as well.

The increase in aggregate shock λ increases the variability of the relative price of y. All agents demand a larger z and the futures price must increase to clear the futures market.

A larger μ_a represents a more risk-averse population, that chooses greater long positions and smaller short positions. The equilibrium futures price must therefore increase to clear the market.

The increase in the dispersion of risk aversion λ_a increases both the number of short positions less risk averse-agents demand and the number of long positions more risk-averse agents demand. Because of concavity of $z(a)$, larger λ_a results in excess short positions. To clear the futures market, p_F must decrease.

Futures Price and Expected Spot Price

As Proposition 4 shows, futures price is not an unbiased estimator of the expected spot price, $E(p) = \frac{1}{2}\bar{p} + \frac{1}{2}\underline{p}$. The presence of backwardation or contango, defined here as the relation between the futures price and the expected spot price, can be tied directly to the distribution of the random output of y, and to the risk aversion of the population.[6]

[6] Most of the current literature studies backwardation as an increasing trend and contango as a decreasing trend in the term structure of futures prices (e.g. Kolb, 1992). Originaly, Keynes and Hicks considered normal backwardation to be the equivalent of a risk premium. Keynes (1930, p.144) wrote: " The quoted forward price,..., must fall below the anticipated future spot price by at least the amount of the normal backwardation."

The effect of the average coefficient of risk aversion, μ_a, on the futures price is intuitive. Equilibrium futures price is a weighted average of $\overline{p}$ and $\underline{p}$. If the population consists of strongly risk averse individuals (μ_a is relatively large) who tend to reduce their risk in utility by taking larger long positions, the futures market clears at a futures price that is relatively high, close to $\overline{p}$. This results in contango, the upward bias of the futures price $p_F > E(p)$. On the other hand, if the population consists of less risk averse agents (μ_a is relatively small) who tend to increase their risk in utility by taking large short positions, the equilibrium is achieved at a relatively low futures price, close to $\underline{p}$. The futures price is biased downward and backwardation $p_F < E(p)$ is observed.

The relationship between the futures price and the expected spot price is represented in the following three diagrams. They show the ratio $p_F/E(p)$ as a function of the aggregate shock for three different values of the average coefficient of risk aversion. When backwardation occurs, $p_F/E(p) < 1$, while $p_F/E(p) > 1$ shows contango.

The values $\mu = 1$ and $\lambda_a = 1$ are the same in all three diagrams. In Figure 2, average risk aversion, $\mu_a = 2$, is weak; in Figure 3, $\mu_a = 3$ is moderate; Figure 4 shows strong average risk aversion, $\mu_a = 4$.

Equilibrium With A Binding No-Default Constraint

Even absent the no-default constraint, each agent's choice of the futures position is bounded from above by

$$z = \frac{\left(\widehat{\beta}\left(\mu+\lambda\right)^{1-\beta} - \widehat{\beta}\left(\mu-\lambda\right)^{1-\beta}\right)}{\widehat{\beta}\left(\mu-\lambda\right)^{1-\beta}\overline{D} + \widehat{\beta}\left(\mu+\lambda\right)^{1-\beta}\underline{D}},$$

for all $a \in (\mu_a - \lambda_a, \mu_a + \lambda_a)$. The expression on the right-hand side is a position an extremely risk averse agent ($a \to \infty$) would choose; he would be fully insured, purchasing all of his y in the futures market, none in the spot market. Therefore, the no default constraint never binds for those who take long positions.

There is, however, no corresponding lower bound on the short futures position an agent with low risk aversion might choose. The income of an agent in the low output/high price of y state of the world is

$$I^x = I^y = 1 + \overline{D}z.$$

A short position must be constrained by $z \geq -\frac{1}{\overline{D}}$. An agent with a low risk aversion (a sufficiently small a), would choose $z < -\frac{1}{\overline{D}}$, making his income negative in the low-output state. When the constraint is binding for the least risk averse agents, properties of the equilibrium described in parts A) and B) of Proposition 4 do not change. The futures price, however, must be greater than the value given by the formula in part C) of Proposition 4.

Next section shows that the properties of equilibrium are different when the output of good y is subject to a combination of the aggregate and idiosyncratic shocks.

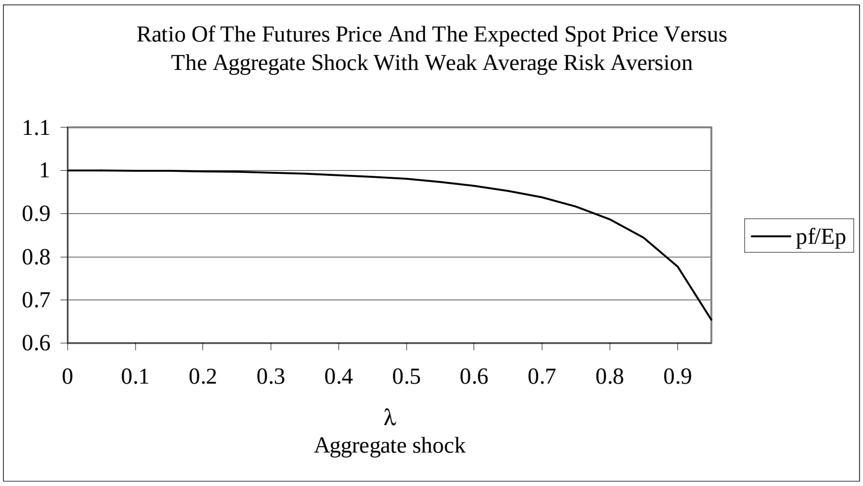

Figure 2: Weak average risk aversion ($\mu_a = 2$) results in backwardation for all λ

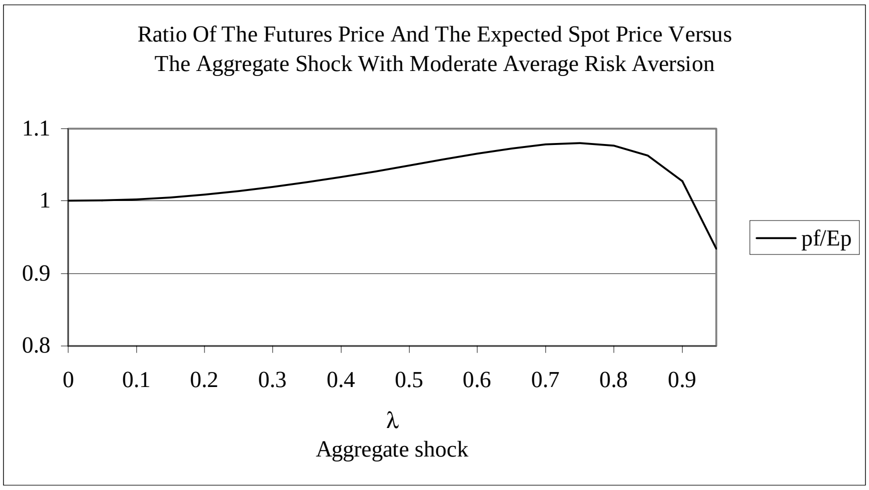

Figure 3: Moderate average risk aversion ($\mu_a = 3$) results in backwardation for large λ, contango for small λ

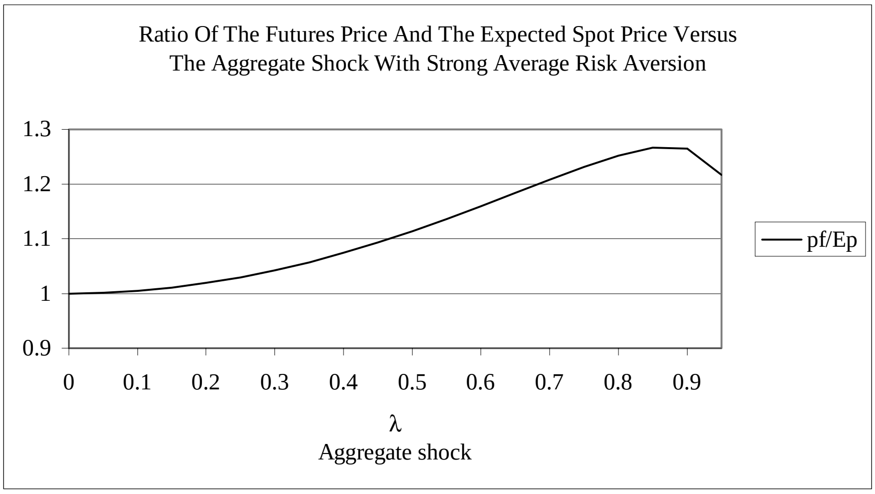

Figure 4: Strong average risk aversion ($\mu_a = 4$) results in contango for all λ

3 Model With Aggregate And Idiosyncratic Shocks

Commodity producers are often exposed to output risk that is not related to the overall market conditions. For example, even when the country enjoys a bumper crop, an individual farmer may loose his crop to a localized flood. In this section, the effects of such idiosyncratic risk on agents' behavior and market equilibrium are examined.

3.1 Industry Formation Without A Futures Market

This section considers the effects of a combination of aggregate and idiosyncratic shocks on equilibrium, when there is no futures market. The idiosyncratic shock is modeled as a Bernoulli random variable $\tilde{\omega}_a$ that has realizations $\omega > 0$ with probability $\frac{1}{2}$ and $-\omega$ with probability $\frac{1}{2}$. While all y-producers experience the same realization of the aggregate shock, the realization of the idiosyncratic shock differs among y-producers. It is assumed that the sum of realizations of the idiosyncratic shock over the subset H of all y-producers is zero: $\int_H \widetilde{\omega}_a d\nu(a) = 0$. There are now four possible realizations of a y-producer's random output.

$$\widetilde{Y} = \begin{cases} \mu + \lambda + \omega, \text{ with probability } 1/4 \\ \mu + \lambda - \omega, \text{ with probability } 1/4 \\ \mu - \lambda + \omega, \text{ with probability } 1/4 \\ \mu - \lambda - \omega, \text{ with probability } 1/4 \end{cases}.$$

It is assumed that $0 < \omega \leq \mu - \lambda$ for all y-producers; output is non-negative in the worst state. The expected utility function of an x-producer is identical to (1):

$$EV^x = -\frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\underline{p}^{1-\beta}}\right] - \frac{1}{2}\exp\left[-a\frac{\widehat{\beta}}{\overline{p}^{1-\beta}}\right].$$

The y-producer's expected utility function is

$$\begin{aligned} EV^y &= \frac{1}{2}\left(\frac{1}{2}\exp\left[-a\widehat{\beta}\underline{p}^{\beta}(\mu+\lambda+\omega)\right] + \frac{1}{2}\exp\left[-a\widehat{\beta}\underline{p}^{\beta}(\mu+\lambda-\omega)\right]\right) \\ &\quad + \frac{1}{2}\left(\frac{1}{2}\exp\left[-a\widehat{\beta}\overline{p}^{\beta}(\mu-\lambda+\omega)\right] + \frac{1}{2}\exp\left[-a\widehat{\beta}\overline{p}^{\beta}(\mu-\lambda-\omega)\right]\right). \end{aligned}$$

Expressions containing idiosyncratic shock ω can be factored out as hyperbolic cosine and EV^y can be conveniently written as

$$EV^y = -\frac{1}{2}\underline{C}\exp\left[-a\widehat{\beta}\underline{p}^{\beta}(\mu+\lambda)\right] - \frac{1}{2}\overline{C}\exp\left[-a\widehat{\beta}\overline{p}^{\beta}(\mu-\lambda)\right], \qquad (8)$$

where

$$\begin{aligned} \underline{C} &\equiv \cosh\left(a\widehat{\beta}\underline{p}^{\beta}\omega\right) = \frac{\exp\left[a\widehat{\beta}\underline{p}^{\beta}\omega\right] + \exp\left[-a\widehat{\beta}\underline{p}^{\beta}\omega\right]}{2}, \text{ and} \\ \overline{C} &\equiv \cosh\left(a\widehat{\beta}\overline{p}^{\beta}\omega\right) = \frac{\exp\left[a\widehat{\beta}\overline{p}^{\beta}\omega\right] + \exp\left[-a\widehat{\beta}\overline{p}^{\beta}\omega\right]}{2}. \end{aligned}$$

Market Equilibrium

Quantity demanded of good x is the sum of the quantities demanded by x-producers and y-producers:

$$\beta A + \beta \int_H \widetilde{p}\left(\mu + \widetilde{\lambda} + \widetilde{\omega}_a\right) d\nu(a).$$

Since $\int_H \widetilde{\omega}_a d\nu(a) = 0$, quantity demanded reduces to

$$\beta A + \beta \int_H \widetilde{p}\left(\mu + \widetilde{\lambda}\right) d\nu(a).$$

The supply of good x is equal to A. The market for good x clears when

$$\beta A + \beta(1-A)\underline{p}(\mu+\lambda) = A$$

if the aggregate shock is high or

$$\beta A + \beta(1-A)\overline{p}(\mu-\lambda) = A$$